

**UNCERTAINTIES IN OBTAINING HIGH RELIABILITY
FROM STRESS-STRENGTH MODELS**

502-38

51390

p. 19

Donald M. Neal, William T. Matthews, and Mark G. Vangel

**U.S. Army Materials Technology Laboratory
Watertown, Massachusetts****ABSTRACT**

There has been a recent interest in determining high statistical reliability in risk assessment of aircraft components. This report identifies the potential consequences of incorrectly assuming a particular statistical distribution for stress or strength data used in obtaining the high reliability values. The computation of the reliability is defined as the probability of the strength being greater than the stress over the range of stress values. This method is often referred to as the stress-strength model.

A sensitivity analysis was performed involving a comparison of reliability results in order to evaluate the effects of assuming specific statistical distributions. Both known population distributions, and those that differed slightly from the known, were considered. Results showed substantial differences in reliability estimates even for almost nondetectable differences in the assumed distributions. These differences represent a potential problem in using the stress-strength model for high reliability computations, since in practice it is impossible to ever know the exact (population) distribution.

An alternative reliability computation procedure is examined involving determination of a lower bound on the reliability values using extreme value distributions. This procedure reduces the possibility of obtaining nonconservative reliability estimates. Results indicated the method can provide conservative bounds when computing high reliability.

PRECEDING PAGE BLANK NOT FILMED

503

PAGE 502 INTENTIONALLY BLANK

INTRODUCTION

There has been an interest in quantitative reliability-based structural design for many years. An early example is the structural reliability development by Freudenthal.¹ Stress-strength reliability computations are a principal consideration in structural reliability design. Reliability methods have been considered for many structural applications including: civil engineering,² nuclear reactors,³ fixed wing aircraft,⁴ rotorcraft,⁵ and space vehicle propulsion systems.⁶ Very high structural reliability is expected to be achieved for most applications. A reliability goal of 0.9₍₉₎ per flight hour was suggested in 1955 by Lundberg⁴ for fixed wing civil aircraft. Recently, Lincoln,^{7,8} using reasoning similar to that of Lundberg, cited a reliability goal of 0.9₍₇₎ per flight for fixed wing military aircraft. The U.S. Army has instituted a new structural fatigue integrity criterion for rotorcraft which has been interpreted⁵ as a requirement for a lifetime reliability of 0.9₍₆₎.

The use of advanced materials whose structural properties are best characterized on a statistical basis appears to be a stimulant for increased interest in statistical-based structural design for airborne structures.

A significant feature associated with predictions of structural reliability is that the consequence of a failure event may be more than reduced system performance or the inconvenience of a system being out of service; structural failure can be catastrophic in terms of loss of life and property. In this context it is imperative to evaluate the sensitivity of structural reliability predictions to uncertainties. It appears that this issue has received little attention except for a brief note by Harris and Soms⁹ and a recent presentation by Berens.¹⁰

There are many issues to be faced in obtaining quantitative structural reliability predictions. Such issues include system complexity (many components, multiple failure modes in each component, and interdependence of component behavior), sample or data set size associated with structural loading spectrum conditions and with mechanical properties, and the basis for characterizing structural qualification tests (the number of duplicate specimens and methods for compensation for untested effects such as the effect of environment).

In addition, when predictions of structural behavior are required in the high reliability range, since sufficiently large data sets are usually not available, it is necessary to use parametric modeling methods. Assumed parametric functions permit extrapolation from available data to determine the probability of failure. Since the probability of failure is extremely small, this will always involve substantial extrapolation from what can be observed experimentally. The estimated reliability will therefore depend strongly on the assumed parametric probability density function (PDF). Slight deviation from the assumed model in tail regions can have a dramatic effect on high reliability estimates.

In fact, one might argue, as does Freudenthal,¹¹ that because of the extrapolation involved, statistically-based high reliability calculations for complex systems must always be suspect:

"When dealing with probabilities a clear distinction should be made between conditions arising in design of inexpensive mass products in which the probability figures are derived by statistical interpretation of actual observations or measurements (since a sufficiently large number of observations are actually obtainable), and conditions arising in design of

structures or complex systems. In the latter, probability figures are used simply as a scale or measure of reliability that permits the comparison of alternative designs. The figures can never be checked by observations or measurement since they are obtained by extrapolations so far beyond any possible range of observation that such extrapolation can no longer be based on statistical arguments but could only be justified by relevant physical reasoning. Under these conditions the absolute probability figures have no real significance"

Nonparametric stress-strength procedures do not require specific parametric assumptions, and so it might be hoped that such procedures could circumvent this difficulty. However, Johnson¹² has noted that "The nonparametric approach has one serious drawback. In return for its distribution free property, it is not possible to establish high reliability even with moderate sample sizes." With respect to the use of parametric models, Box¹³ has observed "all models are wrong, but some are useful," meaning that no parametric statistical model should be accepted uncritically. Whenever a model is used, it is the obligation of the analyst to investigate the consequences of departures from an assumed model which, though small, are consistent with available data. Harris and Soms⁹ has illustrated a "serious problem in the use of stress-strength relationships in estimating reliability." In particular, "stress-strength models in reliability theory are highly sensitive to small perturbations in extreme tails." The perturbations considered may arise from an alternative mode of failure such as the presence of a flaw in a structure. Further, they note that the problem cannot be eliminated unless "astronomically large sample sizes are employed."

In the following, the examination of the sensitivity of structural reliability estimates focuses attention on one of the previously cited issues: the selection of a parametric PDF. The examination of the sensitivity of stress-strength reliability estimates is extended to additional perturbation effects. The sensitivity of reliability estimates to the selection of parametric models is considered with emphasis on graphical representations. The results are evaluated with regard to the usefulness of parametric stress-strength models for application to the high reliability regime of 0.9₍₆₎ to 0.9₍₇₎ when the consequence of failure may be catastrophic. An alternative reliability computation procedure is examined involving determination of a lower bound on reliability which can be obtained independently of the assumed PDFs.

STRESS-STRENGTH MODEL

The statistical reliability as referred to in this report is determined in the following manner. Shown in Figure 1 is the stress-strength model where $f_2(s)$ and $f_1(S)$ represent the PDFs for the applied stress s and material strength S .

Since the joint probability dR for the strength being greater than s_1 can be written as,

$$dR = f_2(s_1) ds \int_{s_1}^{\infty} f_1(S) dS \quad (1)$$

then the reliability for all s values is

$$R = \int_{-\infty}^{\infty} f_2(s) \left[\int_s^{\infty} f_1(S) dS \right] ds. \quad (2)$$

Normal stress-strength model

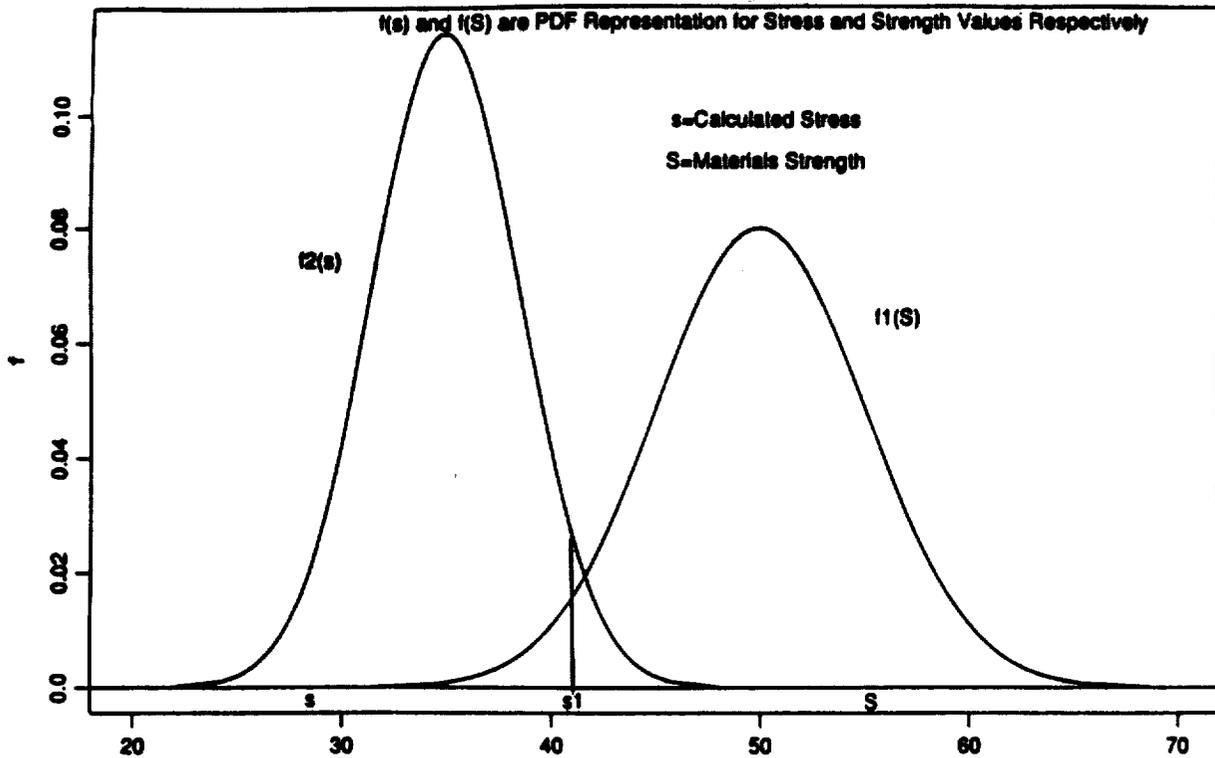


Figure 1. Normal-normal stress-strength model.

PROBABILITY DENSITY FUNCTIONS

A wide variety of PDFs may be applied in obtaining R values. Some examples of PDFs are as follows:

The PDF most often used in stress-strength models is the normal distribution (see Figure 1),

$$f_N(S) = N(S, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{S - \mu}{\sigma} \right)^2 \right\}, \quad (3)$$

where $-\infty < S < \infty$, $\mu > 0$, and $\sigma > 0$. The mean of the population is μ , and the standard deviation σ for this model.

A model which is more easily justified on physical grounds is the Weibull PDF,

$$f_w(S) = \frac{\beta}{\alpha} \left(\frac{S}{\alpha} \right)^{\beta-1} \exp \left[-\left(\frac{S}{\alpha} \right)^\beta \right], \quad (4)$$

where $S > 0$, $\alpha > 0$, and $\beta > 0$. Despite the relevance of the Weibull distribution¹⁴ to the strength of brittle materials, it is not often used, possibly because it is more difficult computationally to obtain reliability values than with the normal model.

If S follows the Weibull PDF, then $\ln(S)$ will have an extreme value distribution with PDF

$$f_{\min}(S) = \frac{1}{b_1} \exp \left[\frac{S-u_1}{b_1} - \exp \left(\frac{S-u_1}{b_1} \right) \right]. \quad (5)$$

The distribution of $-\ln(s)$ is

$$f_{\max}(s) = \frac{1}{b_2} \exp \left[- \left(\frac{s-u_2}{b_2} \right) - \exp \left(- \left(\frac{s-u_2}{b_2} \right) \right) \right]. \quad (6)$$

Both of the above formulas are referred to as extreme value distributions. The use of extreme value distributions in a stress-strength model is illustrated in Figure 2. The extreme value distribution parameters are related to the Weibull parameters as follows:

$$b = \frac{1}{\beta} \quad \text{and} \quad u = -\log \alpha.$$

In order to obtain the population Weibull shape and scale parameters β and α from the known population mean μ and standard deviation σ , the following approximations are suggested:

$$\beta = 1.27 \mu / \sigma - 0.56 \quad (7)$$

and

$$\alpha = \mu / \Gamma \left(\frac{1}{\beta} + 1 \right).$$

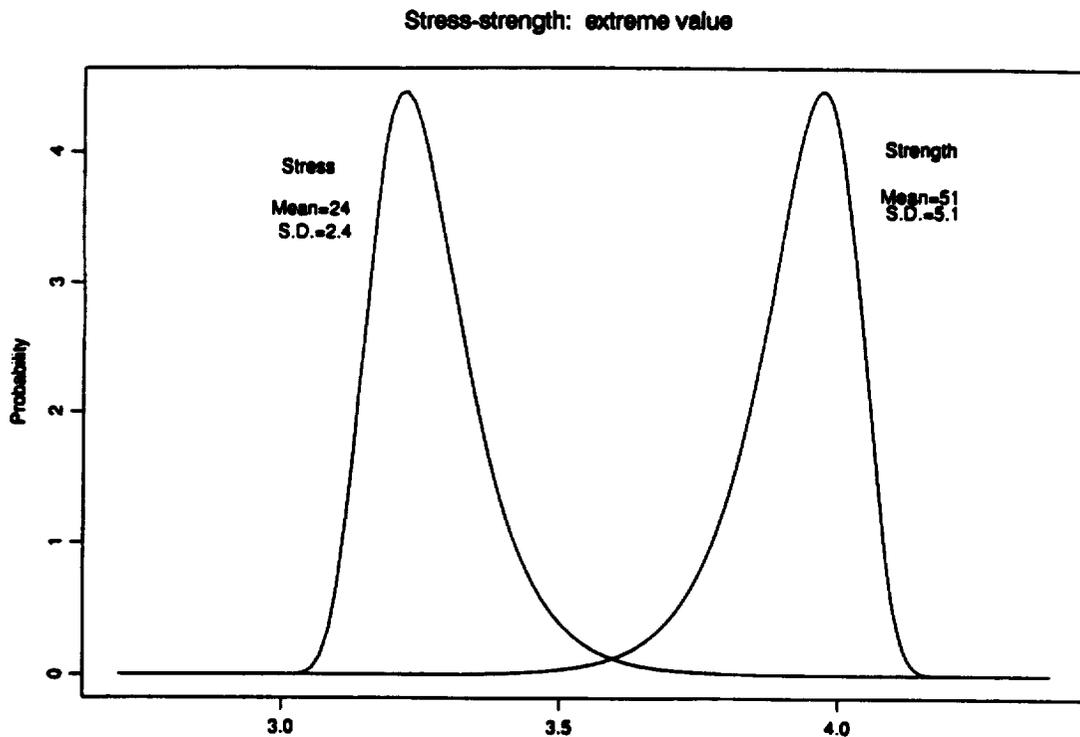


Figure 2. Stress-strength extreme value functions.

The functions defined in Equations 3, 4, 5, and 6 clearly have different shapes and they exhibit dramatically different tail behavior. Since reliability estimates depend strongly on the extreme upper tail of the stress PDF and the extreme lower tail of the strength PDF, the choice of model will typically have a substantial effect on the reliability estimate. For example, R is usually higher when calculated from the normal distribution than when the extreme value model is assumed.

Applying PDFs that are capable of obtaining accurate high reliability estimates (e.g., 0.9₍₆₎) requires prior knowledge of the functional form of the population PDF in addition to the availability of large data sets (e.g., 1,000 replicate specimens). For lower reliability values (e.g., 0.9), a goodness-of-fit test for PDF identification with a moderate amount of data is generally adequate. The consequence of incorrect PDF selection and limited sample sizes are discussed later in this report.

METHODS FOR COMPUTING RELIABILITY R

In determining R from Equation 2 it should be noted that the integration process does not determine an area. The area A described by the intersecting functions in Figure 3 does not represent a 1 - R failure probability. The area A is the probability (P) that either $S < T$ or $s > T$, that is,

$$A = P(S < T) + P(s > T), \quad (8)$$

where T is the point of intersection of the two functions. The area A is obviously not the same as the 1 - R from Equation 2 which determines $P(S > s)$ jointly with $P(s)$.

Stress-strength reliability model: normal-normal

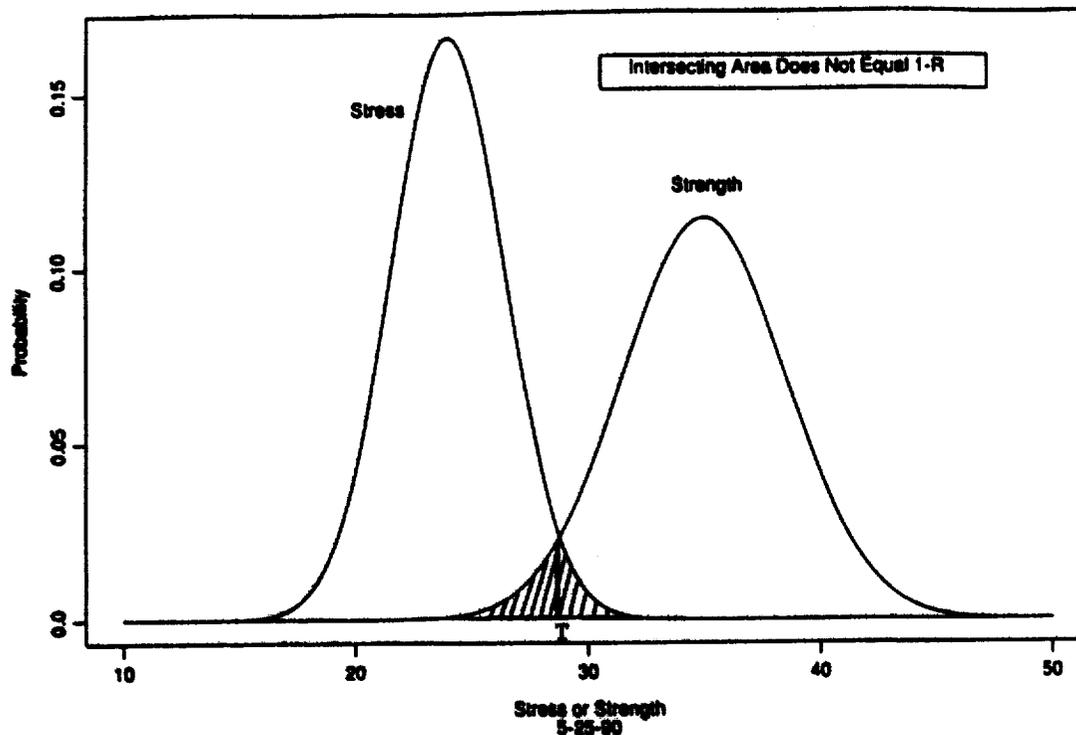


Figure 3. Stress-strength incorrect unreliability region.

Numerical Integration

Numerical integration procedures are usually suggested if a closed form solution of Equation 2 is not available. The numerical integration process involves repeated application of a method such as Simpson's Rule. The inner integral in Equation 2 is evaluated numerically for each ordered s_i value $i = 1, \dots, n$ resulting in an $I_1(i)$ array of values. Each of $I_1(i)$ is then multiplied by the corresponding $f_2(s_i)$ forming another array $I_2(i) = f_2(s_i)I_1(i)$. R is obtained from the I_2 array by reapplying the numerical integration method. This process will usually provide accurate results for $51 \leq n \leq 101$, where n is the number of mesh points in the integration process. Simulation results showed that the limits of integration can be obtained from \pm six standard deviations from the mean.

Closed form solutions are available when the assumed stress and strength PDFs are both normal or both Weibull.

R Computation from Closed Form Solution

If both stress and strength data can be represented by normal PDFs $N(\mu_s, \sigma_s^2)$ and $N(\mu_S, \sigma_S^2)$, respectively, then,

$$R = P(S > s) = \Phi \left(\frac{\mu_S - \mu_s}{\sqrt{\sigma_S^2 + \sigma_s^2}} \right), \quad (9)$$

where (Φ) is the standard $N(0, 1)$ normal cumulative distribution function, μ_s and μ_S are means, and σ_s^2 and σ_S^2 are variances of the stress and strength, respectively.

If both f_1 and f_2 in Equation 2 are Weibull with different scale parameters α_1 and α_2 , but with a common shape parameter β , then the integration indicated in Equation 2 gives the following closed form expression¹²

$$R = \frac{1}{1 + \left(\frac{\alpha_1}{\alpha_2}\right)^\beta} \quad (10)$$

The common shape parameter means that both the stress and strength are skewed in the same way, which is a serious limitation. It is much more reasonable to have a stress distribution with a heavy upper tail and a strength distribution with a heavy lower tail, but this is not possible unless the shape parameters can be varied separately.

Nonparametric Method

This method does not assume a PDF for either stress or strength data. It determines reliability from the ordered array of m stress (s) and n strength (S) values, where each of the S values are compared with all s values. R is the proportion of times $S > s$ for the total number of comparisons, that is

$$R = \frac{1}{m n} \sum_{j=1}^m \sum_{i=1}^n \alpha_i, \text{ where } \alpha_i = \begin{cases} 1, & S_i > s_j \\ 0, & s_j \geq S_i \end{cases} \quad (11)$$

This method is not useful for obtaining high reliability even for relatively large data sets. It is obvious from Equation 11 that for high reliability calculations, mn must be very large; for example, 10^6 would be required in order to obtain R of 0.9(6).

The Weibull, normal, or other parametric PDFs can provide estimates of high R values because of their ability to extrapolate beyond the available empirical data. Unfortunately, the amount of extrapolation dependency determines the magnitude of relative error in R .

CONTAMINATED PROBABILITY DENSITY FUNCTIONS

In order to illustrate the sensitivity of high reliability calculations to small deviations from assumed models, we will take the following approach. Consider the situation where with a high probability of $1 - \epsilon$, specimens are obtained from a primary PDF, while with probability ϵ , specimens come from a secondary PDF. This probability model is referred to as a contaminated model. The secondary component is called the contamination, and the probability ϵ is the amount of contamination.

An example may help clarify this idea. Consider the situation where 97% of the time a specimen is obtained from a population of "good" specimens while the remaining 3% of the time consistently lower strength measurements are obtained, either due to manufacturing defects or to faulty testing. The primary PDF would correspond to the "good" specimens, the contamination would represent the distribution of flawed specimens, and the amount of contamination would be $\epsilon = 0.03$.

The following procedure is introduced in order to examine the effects of computing high reliability values when uncertainties exist in selecting the functions for the stress-strength model. Initially, high reliability values are obtained from the normal stress-strength model (see Equation 9) using known PDFs with different mean values but equal coefficients of variation. The difference in mean values was determined from the required level of high reliability. Another R value is then obtained by applying this known distribution with a small amount of contamination (ϵ) in order to show an almost undetectable difference graphically between the true and contaminated PDFs. The effects of this difference in the reliability computation are discussed in the following sections in order to examine the sensitivity of the stress-strength model to the assumed PDFs. This procedure provides an effective way of demonstrating the effects of assuming a specific PDF in determining high reliability.

The normal PDF with variance contamination for the strength data is,

$$N_{S_v}(\mu_S, \sigma_v^2) = (1 - \epsilon) N(\mu_S, \sigma_S^2) + \epsilon N(\mu_S, K_1 \sigma_S^2), \quad (12)$$

where μ_S and σ_S^2 are the mean and variance for the uncontaminated normal strength distribution, K_1 is a scaling factor, and 100ϵ is the percent contamination.

The strength distribution with location contamination is

$$N_{S_L}(\mu_L, \sigma_S^2) = (1 - \epsilon) N(\mu_S, \sigma_S^2) + \epsilon N(\mu_S \pm K_2 \sigma_S, \sigma_S^2), \quad (13)$$

where K_2 is a scaling factor for the mean μ_S , and the sign determines which tail of the distribution is to be contaminated and σ_S^2 is the variance on $\mu_S \pm K_2\sigma_S$. The location contaminated PDF (see Equation 13) can provide reliability estimates to represent the potential of a secondary failure mode. Contamination of the stress distribution would be similar to that in Equations 12 and 13. It was not necessary to include contaminated distributions for both stress and strength in order to show substantial reduction in the high reliability estimates. The strength PDF contamination was sufficient.

A linear relationship to obtain R for the reliability models when a combination of both contaminated and uncontaminated stress and strength normal PDFs can be written as,

$$R = (1 - \varepsilon_1) (1 - \varepsilon_2) R_{00} + \varepsilon_1 (1 - \varepsilon_2) R_{10} + \varepsilon_2 (1 - \varepsilon_1) R_{01} + \varepsilon_1 \varepsilon_2 R_{11} \quad (14)$$

where $100 \varepsilon_1$ and $100 \varepsilon_2$ are percent contamination for the stress and strength distribution, and the R_{ij} values are obtained for the case of variance contamination only; that is,

$$R_{ij} = \Phi \left(\frac{\mu_S - \mu_s}{\sqrt{\sigma_{S_j}^2 + \sigma_{s_i}^2}} \right) \quad (15)$$

and for location contamination, R_{KL} would be

$$R_{KL} = \Phi \left(\frac{\mu_{SL} - \mu_{sK}}{\sqrt{\sigma_S^2 + \sigma_s^2}} \right) \quad (16)$$

Equation 14 can be extended to include all combinations of variance and location contamination simultaneously, but it was not necessary for this sensitivity analysis. In Equation 16, if $i, j = 0$, then there is no contamination; for $i, j = 1$, then both stress and strength are contaminated. For example, if there is contamination of variance of strength only, then

$$R = (1 - \varepsilon_2) R_{00} + \varepsilon_2 R_{01} \quad (17)$$

where

$$R_{00} = \Phi \left(\frac{\mu_{S_0} - \mu_{s_0}}{\sqrt{\sigma_{S_0}^2 + \sigma_{s_0}^2}} \right)$$

and

$$R_{01} = \Phi \left(\frac{\mu_{S_0} - \mu_{s_0}}{\sqrt{\sigma_{S_1}^2 + \sigma_{s_0}^2}} \right)$$

LOWER RELIABILITY BOUND

A conservative lower bound on the reliability is introduced in order to protect against incorrectly identifying statistical functions in determining high R. The bound is obtained from a method proposed by Bolotin,¹⁵ and modified to employ the extreme value PDFs

(see Equations 5 and 6). The method provides more conservative bounds than would be obtained from standard methods which are dependent on the assumed PDFs. The selection of the extreme value functions provides additional conservatism because of their heavier tails. The method is simple to use and is not restricted to any specified PDF. The reliability bounds are (see Figure 4),

$$1 - W_1 W_2 > R > (1 - W_1)(1 - W_2) \quad (18)$$

where $[(1 - W_1)(1 - W_2)]$ represents the probability $s < s_1$ and $S > S_1$, which can be a somewhat conservative estimate.

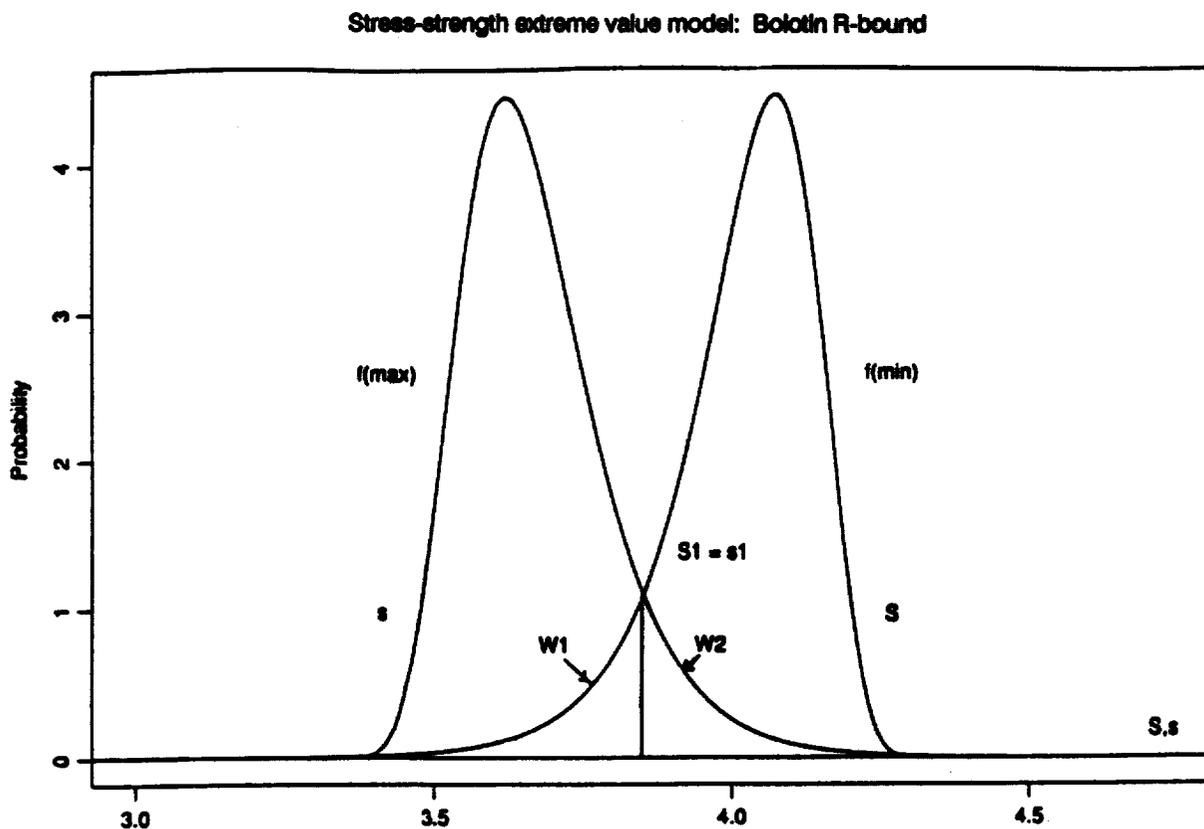


Figure 4. Bolotin reliability bounds using extreme value functions.

The lower bound is then,

$$R_L > (1 - W_1)(1 - W_2) , \quad (19)$$

where

$$W_2 = \int_{s_1}^{\infty} f_2(s) ds \text{ and } W_1 = \int_{\infty}^{S_1} f_1(S) dS$$

for any choice of $s_1 = S_1$.

GOODNESS-OF-FIT TEST

The capability of determining desired PDFs from empirical data was investigated. The choice of PDF will be shown in the following sections to have a substantial effect on high reliability computations, so it is important to examine model selection procedures. A statistical test¹⁶ of goodness-of-fit was introduced in addition to graphical displays in order to select the desired PDFs. Empirical data used in the investigation was obtained by randomly selecting a relatively large number of values from a known normal PDF. A comparison of known contaminated PDFs and the uncontaminated PDF is made with respect to the empirical values.

RESULTS AND DISCUSSIONS

Variance Contamination

Shown in Figure 5a are reliability computation results and a graphical display of a normal/normal stress-strength reliability model, where a 1% ($\epsilon = 0.01$) variance contamination was introduced and scaled by $K_1 = 4$. The graphical display was obtained from application of Equations 12 and 13, where $N(\mu_S, \sigma_S^2)$ is defined in Equation 3. The graph shows an almost undetectable difference between the contaminated and uncontaminated PDFs. This indicates that the choice of ϵ and K are reasonable with respect to the potential differences between assumed and actual PDFs. However, the reliability values differ substantially (0.9₍₆₎ versus 0.998989). This implies that either one failure in a million or 1011 failures in a million is predicted depending on the selection of PDFs which can differ in probability values by less than 0.0005 in the extreme tail regions (see Figure 5b). Using "good" representative PDFs in the stress-strength model in predicting only a single failure will occur in one million operations (e.g., number of flight hours) for $R = 0.9_{(6)}$ can result in a severe anticonservative estimate since for almost identical PDFs, 1011 failures per million could also be predicted.

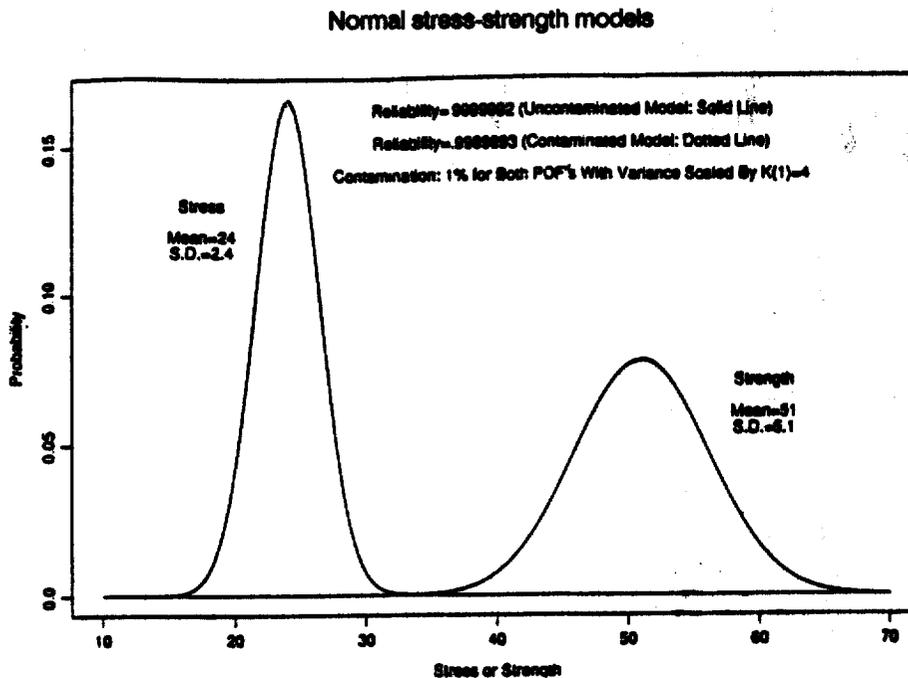


Figure 5a. Stress-strength normal functions with and without variance contamination.

Normal stress-strength models with contamination

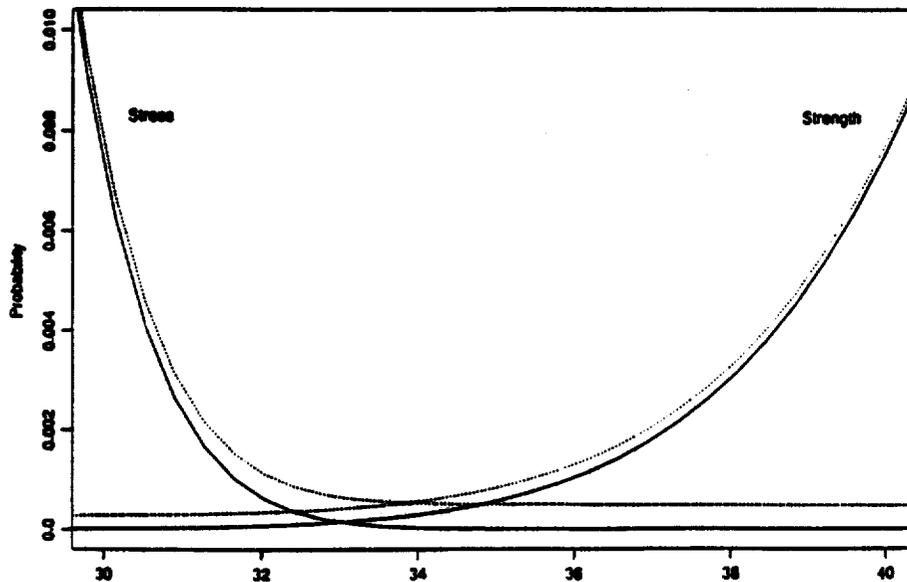


Figure 5b. Detailed region of intersecting functions (see Figure 5a).

The accuracy of the high R estimates depends on the level of precision in defining the extreme tail of PDFs. This requires selecting a PDF from a data set that accurately represents the known population function in the extreme tail regions with a probability difference of much less than 0.0005. Unfortunately, this would require an unrealistically large data set. In current practice, if a very large data set is not available, then PDFs are selected from smaller sets with reliance on the functional representation in regions less than first ordered or greater than the largest value.

The stress-strength procedure is quite effective for the range of R values between 0.5 and 0.95 since usually in the extrapolation process, a small difference in the extreme tail probabilities values will not effect the required accuracy in R. Reliability results from uncontaminated and variance contaminated ($\epsilon = 0.05$ and $K_1 = 5$) PDFs showed no differences for a known $R = 0.95$. Unfortunately, in order to obtain high reliability, extrapolation into the extreme tail of the PDFs is required, thereby increasing the required level of precision necessary to distinguish between, for example, 0.998 and 0.9₍₆₎.

In order to demonstrate the uncertainties in selecting specific PDFs from empirical data when computing high reliability values, the following displays are shown in Figure 6. In Figure 6a, a plot is shown of the empirical normal cumulative density function (CDF) and the corresponding contaminated and uncontaminated normal distribution functions where the mean is 50 and standard deviation (SD) is 5, with sample size $\mu = 100$. Reliability values are also tabulated from the stress-strength model results using all six candidate functions. For example, $R(3, 5)$ is the reliability obtained from variance contamination of 3% and a scale of 5 for variance. A statistical goodness-of-fit test¹⁶ that measures the relative differences in the tail region of the distributions was applied in addition to visual inspection in order to establish if each function could represent the CDF of the ranked data. Results showed this to be true; see Figure 6b for the tabulated observed significance level (OSL) which shows in all cases $OSL > 0.05$, a requirement for the assumed function to be considered from the same population as the empirical data.

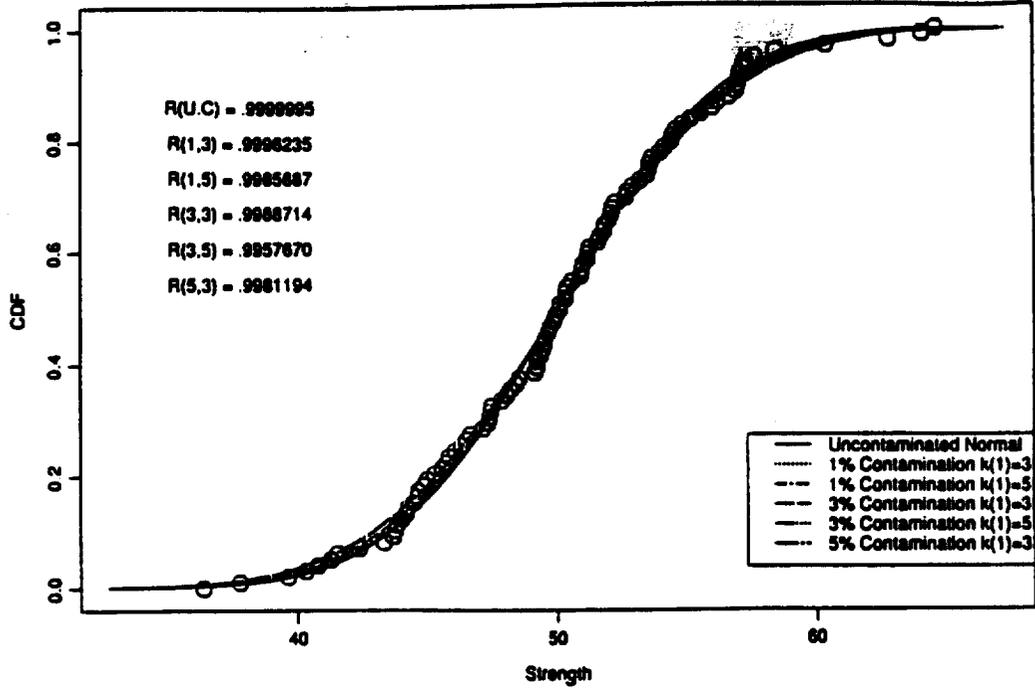


Figure 6a. Goodness-of-fit: empirical versus functional normal CDFs.

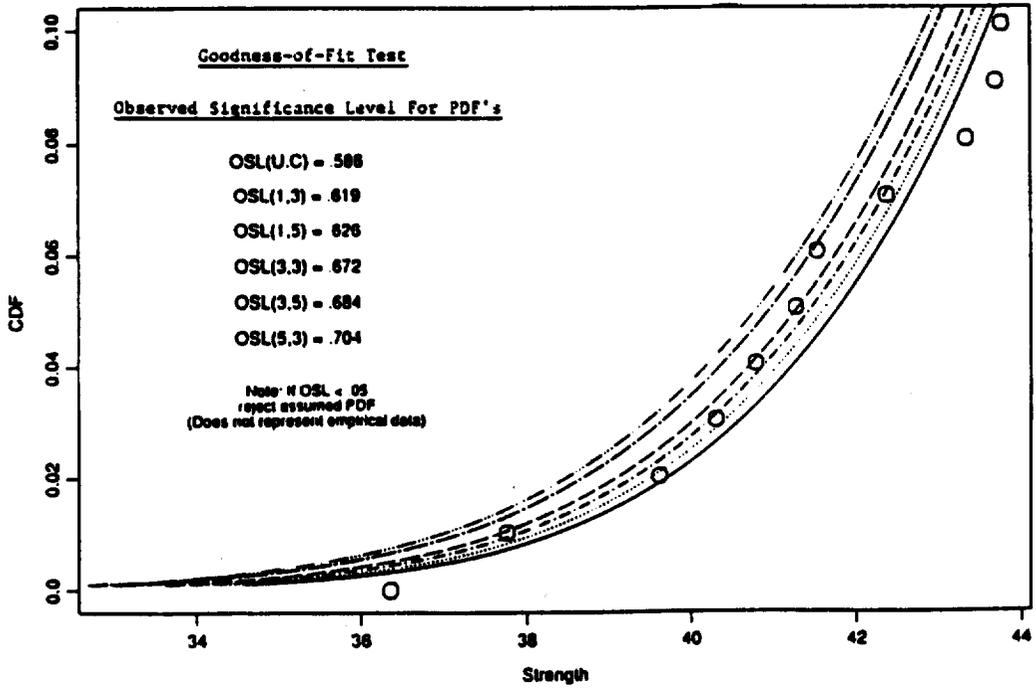


Figure 6b. Lower tail of empirical/normal distributions (see Figure 6a).

The results show that although each distribution fits the data quite well (see Figure 6b), there is a large relative difference in R values: $0.9_{(6)}$ for $R(U.C)$ and 0.9957 for $R(3, 5)$. In Figure 7, the results are similar to those in Figure 5. The variance contamination was 1% with a scale factor of 6 for both σ_S and σ_s . Again, although the functions are similar, the relative reliabilities differ substantially ($0.9_{(6)}$ versus 0.9977197). As was the case in Figure 5, severe consequences could exist if $R = 0.9_{(6)}$ is assumed and the actual reliability was 0.9977197. This could result in a number of premature failures, 2280 in one million, compared to the assumed one failure in a million. The results showed a low level of sensitivity to the selection of the factor K_1 .

Normal stress-strength models

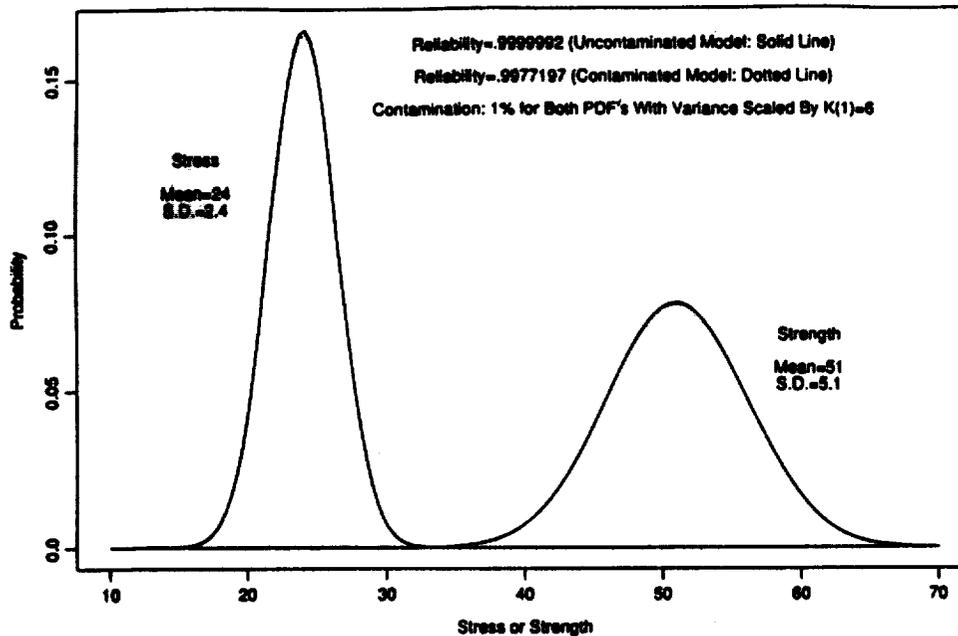


Figure 7. Reliability/normal functions with and without variance contamination.

Location Contamination

In Figure 8, reliability computation results and a graphical display of the stress-strength models are shown. The contaminated functions were obtained from 1% ($\epsilon = 0.01$) location contamination as defined in Equation 13 where $K_2 = 4$ and the (-) value is used for strength and (+) value for stress. The contamination in this case represents a secondary failure mode not considered when assuming a specified function from the test results. For example, ignoring the possibility that one in every 100 parts may have a lower strength level, say 4 standard deviations from the mean, can result in the reliabilities tabulated in the figure. That is, for the assumed correct model, $R = 0.9_{(6)}$, and the actual case where there was a lower strength level having one chance in 100 of occurring resulted in $R = 0.999459$. Figure 9 provides similar results to those in Figure 8 except there is a greater difference in reliability values $0.9_{(6)}$ versus 0.991012 due to a greater shift ($K_2 = 6$) in the mean value for the contaminated PDF. With a 1% contamination this result is predictable since one in a hundred times a failure should occur because $\mu_S - K_2\sigma$ is less than the mean of stress value. The above figure shows the consequences of not being able to identify the correct function because of the inability to always detect a flawed component. The result is the determination of an overly optimistic reliability value when the true reliability could actually be orders of magnitude less.

Normal stress-strength models

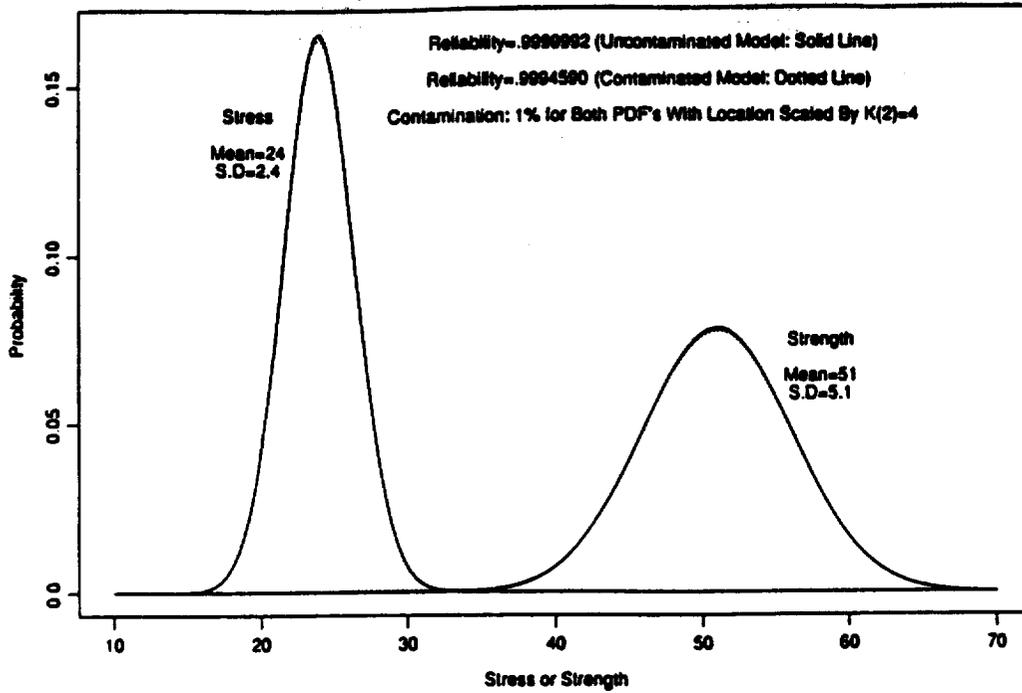


Figure 8. Reliability/normal functions with and without location contamination.

Normal stress-strength models

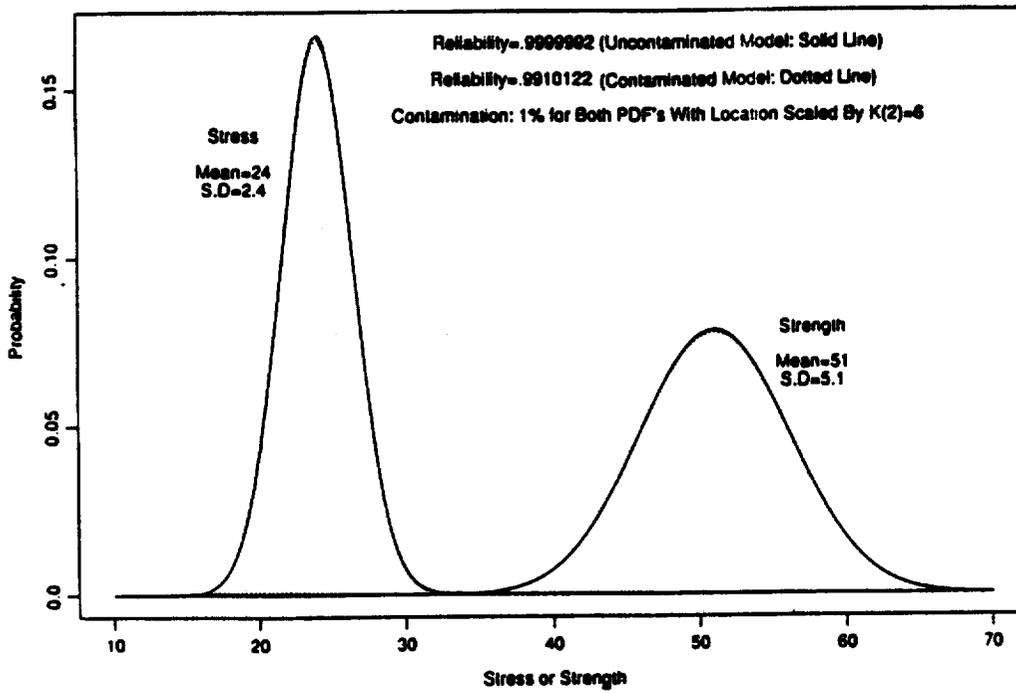


Figure 9. Reliability/normal functions with and without location contamination.

The results in Figure 10 are similar to those in Figure 5 except these were obtained from $\epsilon = 0.03$ and $K_1 = 3$. If the estimated $R = 0.9987350$ is obtained from the empirical data and a higher R value is required ($R = 0.9_{(6)}$), then a material with either greater strength or less contamination would be required. In order to obtain the required $0.9_{(6)}$ from the original contaminated model, a mean strength of 87 is required (see Figure 11). The mean of 87 requirement may not be acceptable to the designer, but this situation can occur if there is a substantial amount of dispersion in the strength data resulting in a long-tailed PDF. The above situation shows when a potentially over-design situation could occur because of the inability to identify the correct PDF in the stress-strength model due to inherent sensitivity and lack of information in the tail regions. This could prevent a good design from being accepted if it is required that the assessment of the design be based upon reliability only.

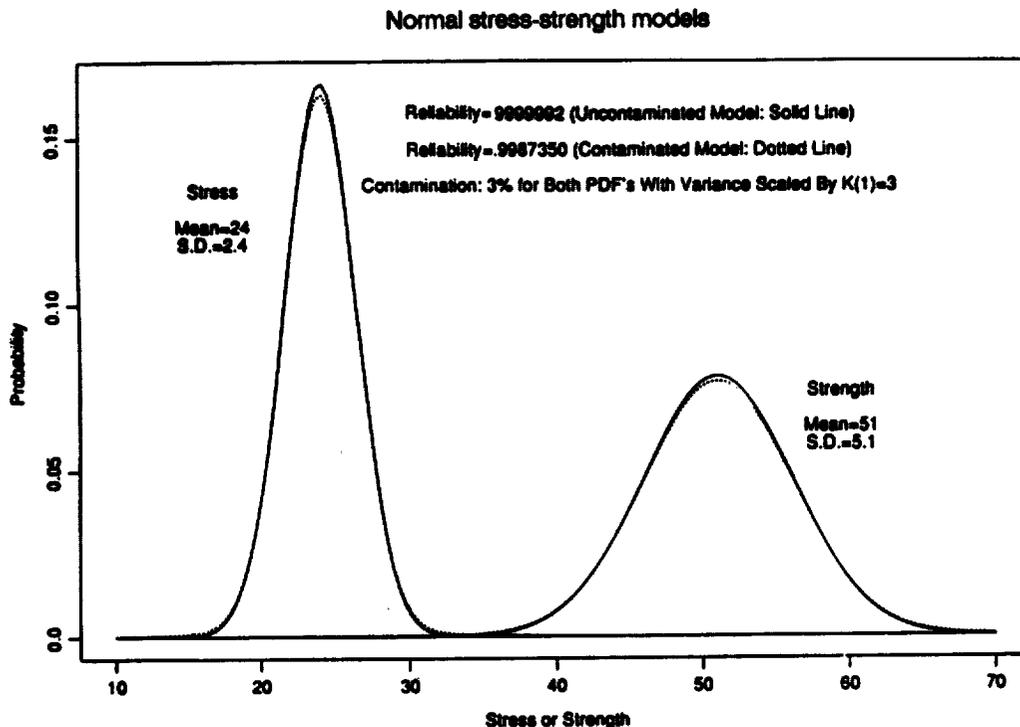


Figure 10. Reliability/stress-strength with and without contamination.

The display in Figure 12 presents four possible reliability values for the case where the means and standard deviation are: stress (24, 2.4) and strength (51, 5.1). The result from the uncontaminated normal is $0.9_{(6)}$. $R = 0.995043$ was obtained from the contaminated PDF application. Since, as was shown previously, the stress-strength model will often provide either relatively very high or low R values depending upon the chance selection of the PDFs. In order to compensate for the uncertainty in selecting the PDFs for stress and strength data, extreme value distributions are introduced (see Figure 2) in the reliability computation. This resulted in $R = 0.999045$. Unfortunately, this did not provide a value lower than the contaminated model result of 0.9950428 . In order to obtain additional conservatism in the R estimate, a modification of a method by Bolotin is examined involving the determination of lower bound on R (see Figure 4 and Equation 19) in conjunction with the extreme value PDFs. The resultant lower bound estimate of 0.9796063 provides a significantly lower value than that of the contaminated model. This was also true for all contaminated models in this study.

Normal stress-strength models

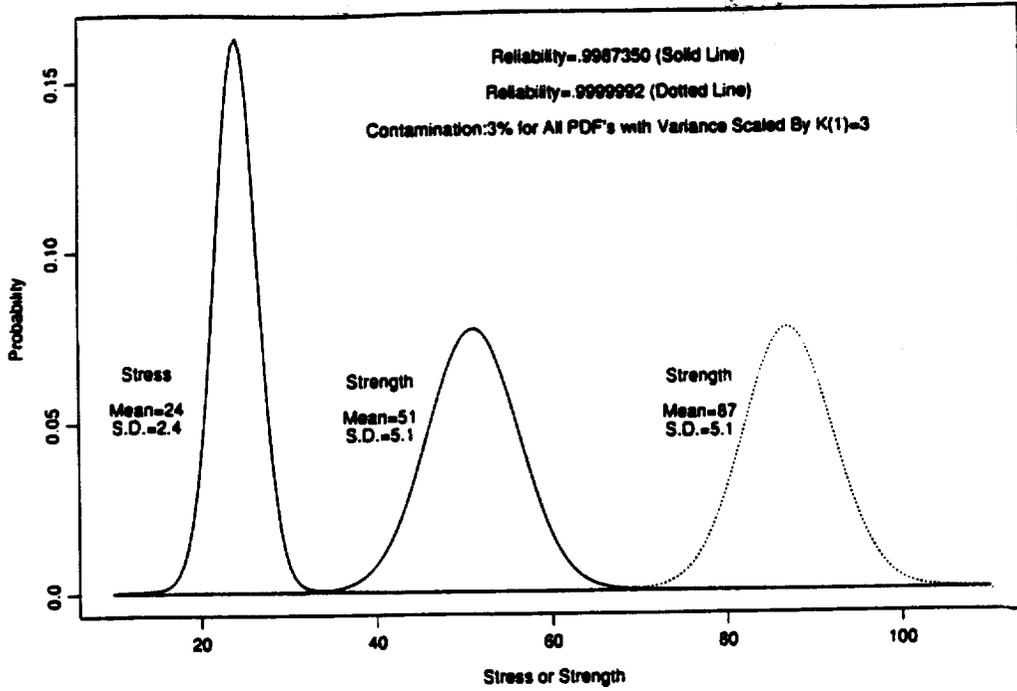


Figure 11. Increased strength requirements for high reliability.

Extreme value stress-strength model

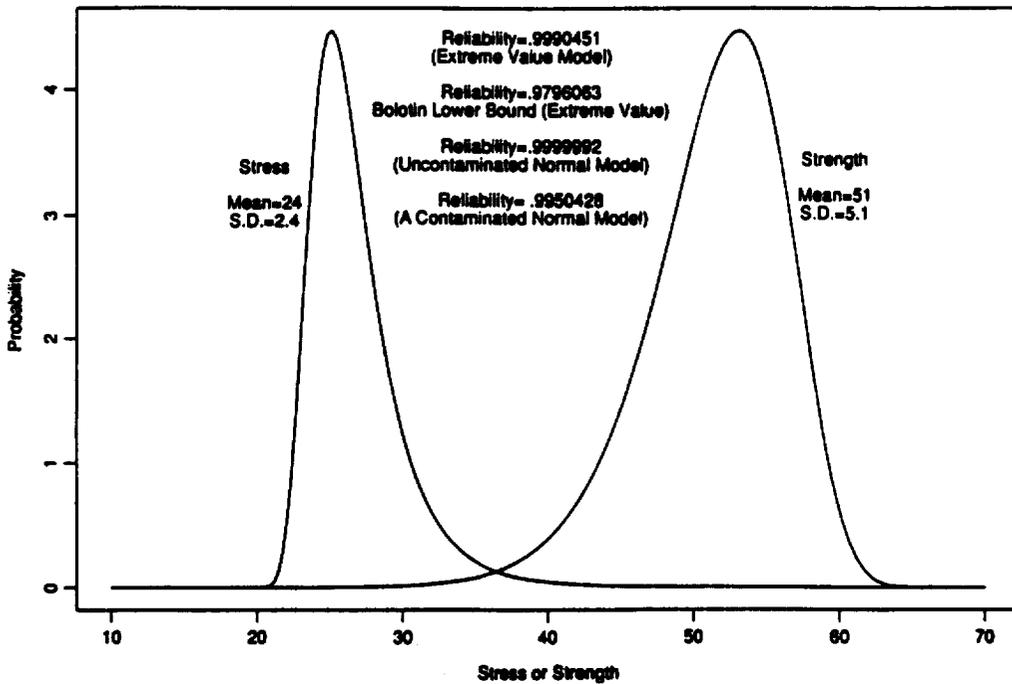


Figure 12. Reliability comparison: PDFs and lower bound.

This lower bound estimate could provide some security in estimating R, although results may be excessively conservative for some practical applications. In Table 1, the distribution of R values as a function of the sample size is presented. R values were obtained from repeated application of the uncontaminated stress-strength model of Figure 5 using randomly selected, normally distributed samples. For a sample size of 10, R ranges from 0.9₍₆₎ to 0.998417 indicating the instability associated with small samples. Higher order quantiles (e.g., 60%) were not included since they were all greater than 0.9₍₆₎.

Table 1. DISTRIBUTION OF R VERSUS SAMPLE SIZE

Distribution (%)	Reliability R			
	Sample Size			
	10	50	100	1000
0.1	0.998417 (1583)	0.999932 (68)	0.999990 (20)	0.999998 (2)
1	0.999160 (840)	0.999968 (32)	0.999991 (9)	0.999998 (2)
10	0.999943 (57)	0.999994 (6)	0.999998 (2)	0.999999 (1)
25	0.999994 (6)	0.999998 (2)	0.999999 (1)	0.999999 (1)
50	0.999999 (1)	0.999999 (1)	0.999999 (1)	0.999999 (1)

() Corresponding Number of Failures Per Million

A sample size of 50 or 100 provides reasonable stability, and a sample of 1000 shows essentially no variability. The results from Table 1 show that for a sample of 1000, an estimate of $R = 0.9_{(6)}$ would be acceptable. This is not necessarily correct since results from the table only address the sample size issue which is independent of the uncertainties in the PDF selection process. There are two requirements for obtaining accurate high reliability values from the strength model: large samples ($n > 1000$) and knowledge of the population PDF. Reliability estimates of 0.95 are much less sensitive to the PDF assumption. If there is a secondary failure mode due to occasional undetected poor manufacturing of the material or an unusually large load occurs that is not accounted for in the design process, then unknown lower reliability values ($R < 0.95$) can exist.

CONCLUSIONS

High reliability estimates from application of the statistical stress-strength model can vary substantially even for almost undetectable differences in the assumed stress and strength PDFs. Specifying high R values (e.g., 0.9₍₆₎) for acceptable structural design can result in higher failure rates than anticipated if the assumed PDFs contain shorter tails than actually exist. Over-design situations can also occur when excessively long-tailed PDFs are applied to the stress-strength model. An effective method for identifying this nonrobust behavior involved application of contaminated and uncontaminated PDFs in the determination of reliability values.

A suggested method for obtaining a lower bound on the reliability estimate provided potentially overly conservative results but was effective in determining values that were lower than any of the R values computed for the contaminated models.

The authors' position regarding the computation of high reliability of $0.9_{(6)}$ agrees with Breiman¹⁷ who says "The probability of failure $P_f = 1 \times 10^{-6}$ has an Alice in Wonderland flavor and should be banned from nonfiction literature." It is therefore recommended that if high reliability calculations are absolutely essential, then the results should be subjected to a sensitivity analysis using contaminated distributions. High reliability values are meaningful only when these values are not substantially affected by an amount of contamination (ϵ) consistent with the sample sizes, and a severity of contamination which is identified by engineering judgement.

ACKNOWLEDGMENT

The authors wish to thank Lucy Ohannesian and the IMA-V group of U.S. Army Materials Technology Laboratory for preparing this manuscript for publication.

REFERENCES

1. FREUDENTHAL, A. M. *The Safety of Structures*. Trans. ASCE, v. 112, 1947.
2. CORNELL, C. A. *A Probability-Based Structural Code*. J. Am. Conc. Inst., v. 66, 1969, p. 974-985.
3. U.S. Nuclear Regulatory Commission. *Reactor Safety Study: An Assessment of Accident Risks in U.S. Nuclear Power Plants*. NRC Report WASH-1400, 1975.
4. LUNDBERG, B. *Fatigue Life of Airplane Structures*. J. of the Aeronautical Sciences, v. 22, no. 6, June 1955, p. 349.
5. ARDEN, R. W., and IMMEN, F. H. *U.S. Army Requirements for Fatigue Integrity*. Proceedings of American Helicopter Society National Technical Specialists Meeting on Advanced Rotorcraft Structures, Williamsburg, VA, October 1968.
6. SHIAO, M. C., and CHAMIS, C. C. *Probability of Failure and Risk Assessment of Propulsion Structural Components*. NASA Technical Memorandum 102323, 1969.
7. LINCOLN, J. W. *Risk Assessment for an Aging Military Aircraft*. J. of Aircraft, v. 22, no. 8, August 1985, p. 687.
8. CORNOG, D. O., and LINCOLN, J. W. *Risk Assessment of the F-16 Wing*. Proceedings of the 1988 Structural Integrity Conference, San Antonio, TX, WRDC-TR-89-8071, Wright-Patterson AFB, Ohio, 1989.
9. HARRIS, B., and SOMS, A. P. *A Note on the Difficulty Inherent in Estimating Reliability from Stress-Strength Relationships*. MRC-2123, Mathematics Research Center, U. of Wisconsin, AD-A073637, October 1980.
10. BERENS, A. *Structural Risk Analysis in Aging Aircraft Fleets*. Proceedings of the 1988 Structural Integrity Conference, San Antonio, TX, WRDC-TR-89-8071, Wright-Patterson AFB, Ohio, 1989.
11. FREUDENTHAL, A. M. *Fatigue Sensitivity and Reliability of Mechanical Systems, Especially Aircraft Structures*. WADD Technical Report 61-53, Wright-Patterson AFB, Ohio, 1961.
12. JOHNSON, R. A. *Stress-Strength Models for Reliability*. Handbook for Statistics, Elsevier Science Publishers, New York, v. 7, 1988, p. 27.
13. BOX, G. E. P. *Robustness in the Strategy of Scientific Model Building*. Robustness in Statistics, R. L. Launer, and G. N. Wilkinson, eds., Academic Press, Inc., New York, 1979.
14. BURY, K. V. *Statistical Models in Applied Science*. John Wiley and Sons, Inc. (New York, London, Sydney, Toronto), 1975.
15. BOLOTIN, V. V. *Statistical Methods in Structural Mechanics*. Holden-Day, Inc., San Francisco, CA, J. J. Brandstatter, ed., 1969.
16. ANDERSON, T. W., and DARLING, D. A. *A Test of Goodness-of-fit*. J. Am. Statis. Assoc., v. 49, 1954, p. 765-769.
17. BREIMAN, L., and STONE, C. *Broad Spectrum Estimates and Confidence Intervals for Tail Quantiles*. Technical Report No. 46, Stat. Dept., University of California, Berkeley, CA, 1985. Also to be published in Journal of Statistical Computation and Simulation.

THIS PAGE INTENTIONALLY BLANK